מבני נתונים ומבוא לאלגוריתמים - מטלת מנחה 11

שאלה 1

נתונה השגרה $Selection\_Sort(A, n)$ שמטרתה למיין איברים במערך A בעל n איברים.

א. הרצת האלגוריתם על המערך $[3, 5, 4, 1, 2]$:

$[1, 5, 4, 3, 2]$
$[1, 2, 4, 3, 5]$
$[1, 2, 3, 4, 5]$
$[1, 2, 3, 4, 5]$

ב. <u>שמורה:</u> בסוף האיטרציה ה-k של הלולאה הפנימית, $A[min] = minimum(A[i],..., A[i + k])$.
נוכיח את השמורה באינדוקציה.

<u>בסיס:</u> לפני תחילת האיטרציה הראשונה מתקיים $min = i$ ולכן $A[min] = A[i]$.
אילו $A[i + 1] < A[i]$ אז האלגוריתם יבצע $min \leftarrow (i + 1)$ ולכן $A[min] = A[i + 1] < A[i]$
אחרת, האלגוריתם לא יבצע כלום ויתקיים $A[min] = A[i] \leq A[i + 1]$
בשני המקרים, בסוף האיטרציה הראשונה יתקיים $A[min] = minimum(A[i], A[i + 1])$, כנדרש.

<u>הנחת האינדוקציה:</u> נניח את נכונות השמורה באיטרציה ה-k של הלולאה הפנימית.

<u>הוכחה:</u> נוכיח את נכונות השמורה באיטרציה ה$(k + 1)$ של הלולאה הפנימית.
לפני תחילת האיטרציה, לפי ההנחה, מתקיים $A[min] = minimum(A[i],..., A[i + k])$.
אילו $A[i + k + 1] < A[min]$, האלגוריתם יבצע $min \leftarrow (i + k + 1)$ ויתקיים,
לפי טרנזיטיביות, אי-השוויון $A[i + k + 1] = A[min] < minimum(A[i], ..., A[i + k])$,
כלומר: $A[min] = minimum(A[i], ..., A[i + k + 1])$ כנדרש.

אחרת, האלגוריתם לא יבצע דבר ויתקיים אי-השוויון הבא:
$A[min] = minimum(A[i],..., A[i + k]) \leq A[i + k + 1]$
ולכן בהכרח $A[min] = minimum(A[i],..., A[i + k + 1])$ כנדרש.

בתום ביצוע הלולאה הפנימית, לפי נכונות השמורה, $A[min]$ יכיל את הערך המינימלי* מבין הערכים $A[i],..., A[n]$ (זוהי משמעות ה"בחירה" באלגוריתם זה - בוחרים את האיבר הקטן ביותר).

* הערה - עלולים להיות שני איברים במערך, או יותר, בעלי אותו ערך מינימלי. במקרה זה, האלגוריתם בוחר את האיבר בעל האינדקס הנמוך ביותר ובכל מקרה יהיה איבר יחיד שייבחר. ההגדרה של מערך ממוין מאפשרת לנו אי-שוויון חלש בין שני איברים סמוכים וכמו כן שמורת הלולאה בסעיף ג' מתחשבת במקרים אלה.

ג. <u>שמורה:</u> בתום האיטרציה ה-i של הלולאה החיצונית, מתקיים $A[1] \leq ... \leq A[i]$, וכן לכל j כך ש $1 \leq i < j \leq n$ מתקיים $A[i] \leq A[j]$.
נוכיח את השמורה באינדוקציה.

<u>בסיס:</u> במהלך ביצוע האיטרציה הראשונה, לפני ביצוע הוראת ה $swap$, לפי סעיף ב' אנחנו יודעים כי $A[min]$ הוא האיבר המינימלי מבין האיברים $A[1],...,A[n]$ שהם כל איברי המערך. לאחר ביצוע הוראה זו , האיבר המינימלי במערך יהיה $A[1]$.
התנאי הראשון, $A[1] \leq A[1]$ מתקיים באופן טריוויאלי. התנאי השני מתקיים משום ש$A[1]$ הוא האיבר המינימלי במערך, ולכן עבור כל j המקיים $1 < j \leq n$ יתקיים $A[1] \leq A[j]$ כנדרש.

<u>הנחת האינדוקציה:</u> יהי $i \geq 1$ כלשהו. נניח את נכונות השמורה עבורו.

<u>הוכחה:</u> נוכיח את נכונות השמורה עבור האיטרציה ה$(i + 1)$ של הלולאה החיצונית.
לפי סעיף ב', לפני ביצוע הוראת ה$swap$ באיטרציה זו, $A[min]$ הוא האיבר המינימלי מבין האיברים $A[i + 1],...,A[n]$.
לכן, לאחר ביצוע ההחלפה, יתקיים $A[i + 1] \leq A[i + 2],...,A[n]$, כלומר לכל j המקיים $i + 1 < j \leq n$ מתקיים $A[i + 1] \leq A[j]$ כנדרש.

לפי ההנחה, מתקיים $A[1] \leq ... \leq A[i]$, וכמו כן עבור $i + 1$ המקיים $i < i + 1 \leq n$ מתקיים אי-השוויון $A[i] \leq A[i + 1]$. לכן, לפי טרנזיטיביות, מתקיים $A[1] \leq ... \leq A[i + 1]$ כנדרש.

בסיום השגרה, בתום ביצוע הלולאה החיצונית (לאחר האיטרציה ה$(n - 1)$) מתקיים לפי נכונות השמורה $A[1] \leq ... \leq A[n - 1]$, וכמו כן עבור n המקיים $n - 1 < n \leq n$ מתקיים $A[n - 1] \leq A[n]$ לכן לפי טרנזיטיביות מתקיים $A[1] \leq ... \leq A[n]$ והמערך ממוין כנדרש.

ד. <u>שמורה:</u> בסוף האיטרציה הi של הלולאה החיצונית, מתקיים $A[1] \leq ... \leq A[i]$.
אילו השמורה נכונה, אזי בסוף האיטרציה ה $(n - 1)$ יתקיים $A[1] \leq ... \leq A[n - 1]$ אך לא נוכל להסיק את אי-השוויון $A[n - 1] \leq A[n]$ ולכן השמורה אינה מאפשרת להוכיח את נכונות האלגוריתם!

גם אם היינו יכולים להסיק אי-שוויון זה, לא נוכל להוכיח את צעד האינדוקציה:

<u>בסיס:</u> באופן טריוויאלי מתקיים $A[1] \leq A[1]$.

<u>הנחה:</u> יהי $1 \leq i \leq n - 1$ ונניח את נכונות השמורה עבורו.

<u>הוכחה:</u> נוכיח את השמורה בסוף האיטרציה ה$(i + 1)$.
לפי סעיף ב', בסוף האיטרציה מתקיים $A[i + 1] = minimum(A[i + 1],...,A[n])$.
לפי ההנחה מתקיים $A[1] \leq ... \leq A[i]$.
לא נוכל להוכיח את אי-השוויון $A[i] \leq A[i + 1]$ בהתבסס על נתונים אלה ולכן לא ניתן להוכיח את נכונות השמורה.

<u>דוגמה נגדית:</u> ניקח את המערך $[1, 5, 2, 6]$.
במערך זה עבור $i = 2$ מתקיים $A[1] \leq A[2]$ וכן $A[3] = minimum(A[3], A[4])$ כנדרש,
ועם זאת לא מתקיים $A[1] \leq A[2] \leq A[3]$.

שאלה 2

נתונה השגרה $What(n)$ המקבלת מספר טבעי n.

לאחר הרצת השגרה מספר פעמים בעזרת טבלאות מעקב הגעתי למסקנה הבאה:\
א. טענה: לכל n טבעי, $What(n) = \lfloor\sqrt{n}\rfloor$, כלומר השגרה תחזיר את החלק השלם של השורש הריבועי של n.

נוכיח את הטענה בעזרת שמורה מתאימה. נשתמש לשם ספירת האיטרציות במשתנה m, זאת משום שלפני תחילת האיטרציה הראשונה ערך המשתנה הוא 1, ובמהלך כל איטרציה מוסיפים לו 1. לכן, לפני האיטרציה ה-k ערכו של m יהיה k.

ב. <u>שמורה:</u> לפני האיטרציה הm מתקיים $i = m^2$.

<u>אילו השמורה נכונה</u>, אזי לפני כל איטרציה בלולאה i יהיה מספר ריבועי הקטן מ-n, ולאחר היציאה

> מהלולאה (לאחר שהתבצעה m פעמים), $i = m^2$ יהיה המספר הריבועי הקטן ביותר הגדול מ-n. במילים אחרות: $(m-1)^2 \le n < m^2$.
> כלומר: $m - 1 \le \sqrt{n} < m$ ולכן לפי תכונות החלק השלם מתקיים $\lfloor\sqrt{n}\rfloor = m - 1$.
> הערך המוחזר הוא אכן $(m - 1)$ וזוהי תוצאת האלגוריתם.

נוכיח את השמורה באינדוקציה.

<u>בסיס:</u> לפני תחילת האיטרציה הראשונה מתקיים $m = i = 1$ ולכן $i = m^2 = 1^2 = 1$ כנדרש.

<u>הנחה:</u> נניח את נכונות הטענה לפני האיטרציה ה-m.

<u>הוכחה:</u> נוכיח את נכונות הטענה עבור האיטרציה ה$m + 1$.

> לפי ההנחה, לפני האיטרציה ה-m התקיים $i = m^2$.
> במהלך האיטרציה, התבצעה ההוראה $i \leftarrow i + 2m + 1$ (לפני שינוי המשתנה m).
> לכן, הערך שנשמר בi לאחר ביצוע הוראה זו הוא $m^2 + 2m + 1 = (m+1)^2$.
> האלגוריתם לא משנה את ערך המשתנה i לאחר מכן ולכן גם לפני תחילת האיטרציה ה $m + 1$ מתקיים $i = (m+1)^2$ כנדרש.

שאלה 3

א. נגדיר לכל $i \in \{1, 2, 3, 4, 5\}$ את הפונקציה $f_i: N \to N$ כך:
לכל n טבעי,

$$f_i(n) = \begin{cases} 1 & \bmod_5(n) = i \\ n & \bmod_5(n) \neq i \end{cases}$$

טענה: לכל $i, j \in \{1, 2, 3, 4, 5\}$ כך ש $i \neq j$ מתקיים $f_i(n) \neq O(f_j(n))$

הוכחה: יהיו $i, j \in \{1, 2, 3, 4, 5\}$ כך ש $i \neq j$ ונניח בשלילה כי $f_i(n) = O(f_j(n))$.
אז קיימים $c, n_0 > 0$ כך שלכל $n \geq n_0$ טבעי מתקיים $f_i(n) \leq c * f_j(n)$.

נבחר את המספר הטבעי $n' = 5\lceil\frac{n_0}{5}\rceil + 5\lceil c\rceil + j$.
קל לראות כי $n' \bmod 5 = j \neq i$. לכן, $f_i(n') = n'$ וכמו כן $f_j(n') = 1$.
מהגדרת המספר נובע $n' > 5 * \lceil\frac{n_0}{5}\rceil \geq n_0$ (ובפרט לפי טרנזיטיביות $n' \geq n_0$).
נציב באי-השוויון הנובע מהנחת השלילה ונקבל $n' \leq c * 1$ ולכן $n' \leq c$

אבל מהגדרת n' נובע אי-השוויון $n' > 5\lceil c\rceil \geq 5c$.
מאחר ו$c > 0$ נובע $5c > c$ ולכן לפי טרנזיטיביות $n' > c$.
סתירה! לא יכול להיות ש $n' > c$ אבל גם $n' \leq c$.

ב. יהיו הקבועים $\alpha_1, \alpha_2, \beta_1, \beta_2$ ויהיו הפונקציות $f(n)$, $g(n)$ המוגדרות כך:

$$f(n) = n^{\alpha_1} \cdot (lglgn)^{\beta_1}$$
$$g(n) = (lgn)^{\alpha_2} \cdot (\frac{n}{lglgn})^{\beta_2}$$

נחשב:

$$\frac{f(n)}{g(n)} = \frac{n^{\alpha_1} \cdot (lglgn)^{\beta_1}}{(lgn)^{\alpha_2} \cdot (\frac{n}{lglgn})^{\beta_2}} = \frac{n^{\alpha_1 - \beta_2} \cdot (lglgn)^{\beta_1 + \beta_2}}{(lgn)^{\alpha_2}}$$

נשים לב, לפי עמוד 34 במדריך הלמידה, לעובדות הבאות:

א. $(lglgn)^{\beta_1 + \beta_2} = o(lg^{\alpha_2} n)$

ב. אם $\alpha_1 - \beta_2 > 0$ אז $lg^{\alpha_2} n = o(n^{\alpha_1 - \beta_2})$

לפי עובדות אלו נסיק כי ל $lg^{\alpha_2} n$ ול $(lglgn)^{\beta_1 + \beta_2}$ אין כלל השפעה על הגידול של המנה כאשר $\alpha_1 - \beta_2 > 0$ ובמקרה זה $\lim_{n \to \infty} \frac{f(n)}{g(n)} = \lim_{n \to \infty} n^{\alpha_1 - \beta_2} = \infty$.

כמו כן, נסיק כי כאשר $\alpha_1 - \beta_2 < 0$ (ולכן $\beta_2 - \alpha_1 > 0$), מתקיים $\lim_{n \to \infty} \frac{f(n)}{g(n)} = \lim_{n \to \infty} \frac{(lglgn)^{\beta_1 + \beta_2}}{n^{\beta_2 - \alpha_1} \cdot (lgn)^{\alpha_2}}$

במקרה זה קצב הגדילה של כל אחת מן הפונקציות במכנה גדול מקצב הגדילה של הפונקציה במונה, ולכן מתקיים $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$ במקרה זה.

במקרה השלישי, כאשר $\alpha_1 - \beta_2 = 0$, מתקיים:

$$\frac{f(n)}{g(n)} = \frac{(lglgn)^{\beta_1+\beta_2}}{(lgn)^{\alpha_2}}$$

כאמור, קצב הגדילה של הפונקציה במכנה גדול מקצב הגדילה של הפונקציה במונה, ועל כן במקרה זה מתקיים $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$.

המסקנות:

1. עבור $\alpha_1 = 1,\ \alpha_2 = 2,\ \beta_1 = 3,\ \beta_2 = 4$ מתקיים $\alpha_1 - \beta_2 = -3 < 0$ ולכן, כאמור, $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$, לכן $f(n) = o(g(n))$.
2. עבור $\alpha_1 = 4,\ \alpha_2 = 3,\ \beta_1 = 2,\ \beta_2 = 1$ מתקיים $\alpha_1 - \beta_2 = 3 > 0$ ולכן, כאמור, $\lim_{n\to\infty} \frac{f(n)}{g(n)} = \infty$, לכן $f(n) = \omega(g(n))$.
3. לא ייתכנו $\alpha_1, \alpha_2, \beta_1, \beta_2$ עבורם $f(n) = \Theta(g(n))$ משום שעבור כל זוג מספרים שנבחר ל α_1 ו β_2 יתקיימו אחד משלושת המקרים שצוינו לעיל ($\alpha_1 - \beta_2 < 0$, $\alpha_1 - \beta_2 = 0$, $\alpha_1 - \beta_2 > 0$) ועל כן אחד מהמקרים הבאים: $f(n) = o(g(n))$ או $f(n) = \omega(g(n))$ ובכל מקרה לא $f(n) = \Theta(g(n))$.

שאלה 4

סעיף א

ננתח את סיבוכיות זמן הריצה של כל צעד באלגוריתם שלנו.

ההוראה	סיבוכיות זמן-ריצה
"בודק האם $n = 1$, ואם כן הוא עוצר"	$O(1)$ (ולכן $T(1) = 1$)
"ביצוע n פעמים מיון הכנסה על החצי השמאלי"	בפעם הראשונה - חצי המערך השמאלי לא בהכרח ממוין ולכן סיבוכיות זמן הריצה היא של המקרה הגרוע ביותר $\Theta(n^2)$. בשאר הפעמים, החלק במערך ממוין ולכן מיון ההכנסה מקבל את קלט המקרה הטוב ביותר* ופועל בסיבוכיות $\Theta(n)$ במשך $(n - 1)$ פעמים, ועדיין $\Theta(n^2)$ בסך הכל.
"30 פעמים מיון מיזוג על החצי הימני של המערך"	$30 \cdot \Theta(n \lg n) = \Theta(n \lg n)$. במיון מיזוג העובדה שהמערך ממוין לא משנה** את סדר הגודל של סיבוכיות זמן הריצה.
"קריאה רקורסיבית על החצי השמאלי בלבד"	$T(\frac{n}{2})$

סיבוכיות סך ה"עבודה", לא כולל הקריאה הרקורסיבית, היא $\Theta(1) + \Theta(n^2) + \Theta(n \lg n) = \Theta(n^2)$ ולכן, נוסחת הנסיגה היא $T(n) = T(\frac{n}{2}) + \Theta(n^2)$.

נפתור את הנוסחה לפי משפט האב עבור $a = 1, b = 2, f(n) = \Theta(n^2), \log_b a = \log_2 1 = 0$.

מקרה 1: קיים $\epsilon > 0$ כך ש $f(n) = \Theta(n^2) = O(n^{0-\epsilon}) = O(\frac{1}{n^\epsilon})$. כמובן שהמקרה לא מתקיים משום ש $\Theta(n^2)$ הגדלה לא יכולה להיחסם מלמעלה על ידי הטור $\frac{1}{n^\epsilon}$, המתכנס ל-0.

מקרה 2: $f(n) = \Theta(n^2) = \Theta(n^0) = \Theta(1)$ וכמובן שהמקרה לא מתקיים משום ש $\Theta(n^2)$ הגדלה לא יכולה להיחסם מלמעלה על ידי טור קבוע.

מקרה 3: קיים $\epsilon > 0$ כך ש $f(n) = \Theta(n^2) = \Omega(n^{0+\epsilon})$ והמקרה מתקיים עבור $\epsilon = 2$ - כל טור הוא חסם הדוק על עצמו מלמטה.

תנאי הרגולריות: קיים $c < 1$ כך שעבור n גדול מספיק מתקיים $f(\frac{n}{2}) \leq c \cdot f(n)$

ואכן, $(\frac{n}{2})^2 = \frac{n^2}{4} \leq c \cdot n^2$ עבור $c \geq \frac{1}{4}$. נבחר $c = \frac{1}{4}$ והטענה מתקיימת.

לכן, $T(n) = \Theta(n^2)$.

נימוקים חשובים:
* לפי הדרך שבה מיון הכנסה עובד, לאחר האיטרציה ה-i תת-המערך $A[1]..A[i]$ ממוין. לכן, בהינתן שכל המערך ממוין, מתקיים $A[i + 1] \geq A[i]$ ועל כן הלולאה הפנימית לא תופעל בקטע זה ותת המערך $A[1]..A[i + 1]$ ממוין כהלכה.
** במהלך מיזוג, אנחנו נדרשים לעבור על כל n איברי תתי-המערך ולשבצם במקום המתאים במערך גם תחילה נשבץ תת-מערך שלם אחד ורק לאחר מכן את איברי תת-המערך השני, כמו שיקרה במערך ממוין. על פעולה זו חוזרים $\lg n$ פעמים ומכאן שהסיבוכיות לא משתנה.

סעיף ב

אילו $n\ mod\ 3\ =\ 0$:

ההוראה	סיבוכיות זמן-ריצה
"מיון הכנסה על השליש הראשון של המערך"	$(\frac{n}{3})^2 = \frac{n^2}{9} = \Theta(n^2)$.
"קריאה רקורסיבית פעמיים על השליש השני ופעמיים על השליש השלישי"	$T(\frac{n}{3}) + T(\frac{n}{3}) + T(\frac{n}{3}) + T(\frac{n}{3}) = 4T(\frac{n}{3})$
"ממזג בין שלושת השלישים"	$\Theta(\frac{n}{3} + \frac{n}{3} + \frac{n}{3}) = \Theta(n)$

סך כל ה"עבודה" (לא כולל הקריאות הרקורסיביות) במקרה זה: $\Theta(n^2) + \Theta(n) = \Theta(n^2)$.
נוסחת הנסיגה במקרה זה: $T(n) = 4T(\frac{n}{3}) + \Theta(n^2)$

נפתור את הנוסחה בעזרת משפט האב עבור $a = 4, b = 3, f(n) = \Theta(n^2), \log_b a = \log_3 4 \approx 1.26$

מקרה 1: קיים $\epsilon > 0$ כך ש $f(n) = \Theta(n^2) = O(n^{1.26-\epsilon})$. המקרה לא מתקיים משום שעבור כל ϵ שנבחר, הטור $\frac{n^{1.26-\epsilon}}{n^2} = \frac{1}{n^{0.74+\epsilon}}$ מתכנס לאפס ולכן n^2 חוסם מלמעלה בצורה לא הדוקה את הטור $n^{1.26-\epsilon}$ ולא יכול להיות שהטור $n^{1.26-\epsilon}$ יחסום את הטור n^2 מלמעלה בו זמנית.

מקרה 2: $f(n) = \Theta(n^2) = \Theta(n^{1.26})$ לא מתקיים כי הטור $\frac{n^{1.26}}{n^2} = \frac{1}{n^{0.74}}$ מתכנס לאפס ועל כן n^2 חוסם מלמעלה בצורה לא הדוקה את הטור $n^{1.26}$ ולא יכול להיות שהטור $n^{1.26}$ יחסום את הטור n^2 מלמעלה בו זמנית.

מקרה 3: קיים $\epsilon > 0$ כך ש $f(n) = \Theta(n^2) = \Omega(n^{1.26+\epsilon})$. המקרה מתקיים עבור $\epsilon = 2 - \log_3 4 \approx 0.74$ ואז ברור שהטור n^2 חוסם את עצמו מלמטה.

תנאי הרגולריות: קיים $c < 1$ כך שעבור n גדול מספיק מתקיים $4f(\frac{n}{3}) \le c \cdot f(n)$

ואכן, $4(\frac{n}{3})^2 = \frac{4n^2}{9} \le c \cdot n^2$ עבור $c \ge \frac{4}{9}$. נבחר $c = \frac{4}{9}$ והטענה מתקיימת.

לכן, $T(n) = \Theta(n^2)$ במקרה זה.

אילו $n\ mod\ 3\ =\ 1$:

ההוראה	סיבוכיות זמן-ריצה
"מציאת האיבר המינימלי במערך והצבתו במקום הראשון"	$\Theta(n)$
"קריאה רקורסיבית על כל המערך מלבד האיבר הראשון"	$T(n - 1)$

נוסחת הנסיגה במקרה זה: $T(n) = T(n - 1) + \Theta(n)$.
לאחר שרטוט עץ, הגעתי לסכום:

$$T(n) = \sum_{i=0}^{n}(n - i) = \sum_{i=0}^{n} n - \sum_{i=0}^{n} i = n^2 - \frac{n(n+1)}{2} = \Theta(n^2)$$

אילו $n\ mod\ 3\ =\ 2$:

ההוראה	סיבוכיות זמן-ריצה
"מציאת האיברים המינימלי והמקסימלי במערך והצבתם במקום הראשון והאחרון"	$\Theta(n)$
"קריאה רקורסיבית על כל המערך מלבד האיברים הראשון והאחרון"	$T(n\ -\ 2)$

נוסחת הנסיגה במקרה זה: $T(n)\ =\ T(n\ -\ 2)\ +\ \Theta(n)$
לאחר שרטוט עץ, הגעתי לסכום:

$$T(n)\ =\ \sum_{i=0}^{\frac{n}{2}}(n\ -\ 2i)\ =\ \sum_{i=0}^{\frac{n}{2}}n\ -\ \sum_{i=0}^{\frac{n}{2}}2i\ =\ \frac{n^2}{2}-\frac{\frac{n}{2}(n+1)}{2}=\ \Theta(n^2)$$

קיבלנו שבשלושת המקרים, סיבוכיות זמן הריצה של האלגוריתם היא $\Theta(n^2)$ וכך גם סיבוכיות זמן הריצה עבור המקרה הגרוע ביותר.